

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 80 (2020) 2, 321—334

Besedilo prejeto/Received:08/2020; sprejeto/Accepted:08/2020

UDK/UDC: 004.89:17

DOI: 10.34291/BV2020/02/Strahovnik

© 2020 Strahovnik et al., CC BY 4.0

Vojko Strahovnik, Jonas Miklavčič in Mateja Centa

Etični vidiki uporabe algoritemskega odločanja in ostalih sistemov UI v času pandemij oz. izrednih razmer

Ethical Issues Related to the Use of AI-Based Algorithms in Pandemics and Other Public Health Emergencies

Povzeteke: Netransparentnost delovanja je problem mnogih področij rabe umetne inteligence (UI) – tudi, in morda še posebej, sistemov algoritemskega odločanja. Problem ni le tehnične narave, temveč tudi etične. Kljub temu se algoritemski sistemi, ki temeljijo na umetni inteligenci, uporabljajo vse pogosteje in na vse več področjih. V pomoč so nam tudi v času izrednih razmer in trenutna pandemija COVID-19 ni izjema. V času te pandemije so se mnogi sistemi UI izkazali za izredno uporabne, vseeno pa njihova uporaba – v veliki meri tudi zaradi njihove pogosto netransparentne narave – odpira mnoga etična vprašanja. Postavljeni smo pred dilemo, ali je uporaba uspešnih sistemov, ki pa ne delujejo transparentno, lahko (v izrednih razmerah) etično sprejemljiva. Kot možni se kažejo predvsem trije odzivi: prilagoditev tradicionalnega pojmovanja transparentnosti pri odločanju, prilagoditev normativnih izhodišč ali pa omejitev uporabe netransparentnih sistemov algoritemskega odločanja.

Ključne besede: transparentnost, umetna inteligenca, etika izrednih razmer, pandemija, COVID-19

Abstract: Non-transparency is a problem in many areas of artificial intelligence - including, and perhaps especially, algorithmic decision-making systems. The problem is not only technical but also ethical. Nevertheless, algorithmic systems based on artificial intelligence are being used more and more frequently and in more and more areas. They also help us during emergencies, and the current COVID-19 pandemic is no exception. During this pandemic, many AI systems proved to be extremely useful, yet their use, largely due to their often-non-transparent nature, raises many ethical questions. We are faced with the dilemma of whether the use of successful systems that do not operate transparently can be ethically acceptable (in emergencies). The possible responses are threefold: the adaptation of the traditional notion of transparency in decision-making, the adaptation of normative basis and the restriction of the use of non-transparent algorithmic decision-making systems.

Key words: transparency, artificial intelligence, emergency ethics, pandemic, COVID-19

1. Uvod

Osrednji cilj prispevka je raziskati etične vidike oz. izhodišča za presojo uporabe algoritemskega odločanja in drugih sistemov umetne inteligence (UI) v času pandemij oz. izrednih zdravstvenih razmer. Prispevek se tako umešča v niz splošnejših razprav o etični sprejemljivosti algoritemskega odločanja, pri katerem pogosto umanjka t. i. transparentnost, kakor tudi v okvir razprav o etiki izrednih razmer oziroma nujnih primerov (angl. *ethics of emergencies*). V prispevku najprej opredelimo problem transparentnosti, v tretjem razdelku nato izpostavimo nekatere etične dileme, ki rabo teh tehnologij spremljajo. V četrtem razdelku predstavljamo nekaj primerov uporabe takšnih tehnologij in sistemov odločanja ter jih analiziramo z vidika problematike transparentnosti. V petem, sklepnem razdelku izpostavljamo napetost med problemom transparentnosti ter zahtevami transparentnosti, ki jih vključujejo prevladujoče etike javnega zdravja in nujnih primerov oz. izrednih zdravstvenih razmer. Pokazati skušamo, kako raba obravnavanih tehnologij ni združljiva z izhodišči slednje in na koncu zarisati nekaj možnih razrešitev te napetosti.

Poleg mnogih drugih ved in znanosti je tudi teologija poklicana k temu, da odgovarja na etične izzive sodobnega sveta, ki ga zaznamuje prevlada tehnologije (Petkovšek 2019). Literature in dogodkov (konference, simpoziji), ki razmišljajo o vlogi teologije v sodobnem svetu, je vse več. Profesor moralne teologije in bioetike na Papeški Univerzi Gregoriana v Rimu, br. Paolo Benanti, je na konferenci, ki je na temo etike umetne inteligence potekala konec februarja 2020 v Rimu, pomenljivo dejal: »Sledeč primeru elektrike, umetna inteligenca ni potrebna za izvajanje neke specifične dejavnosti, temveč je namenjena spreminjanju samega načina, kako izvajamo vsakodnevna dejanja.« (Eliazàt 2020) Zaradi razdrobljenosti področij vpliva umetne inteligence (UI) je vpliv težje sledljiv in manj opazen, a morda ravno zaradi svoje ‚skritosti‘ toliko bolj zaznamuje naša življenja – in je tako za etični premislek zelo relevantna tema. Umetna inteligenca s svojim prispevkom in tudi pomisleki, ki so z njo povezani, sega na mnoga področja naših življenj in pandemija COVID-19 pri tem ni izjema – pravzaprav je izvrsten primer. Umetna inteligenca se torej tudi v času pandemije COVID-19 ne uporablja za izvajanje specifične zahtevne naloge, temveč je – kot bomo videli – prisotna pri mnogih korakih, področjih, problemih in rešitvah.

2. Problem transparentnosti algoritemskih sistemov odločanja

Pri umetni inteligenci (UI) in strojnem učenju (SU) se ‚problem transparentnosti‘ (uporabljata se tudi izraza ‚problem motnosti oz. nepreglednosti‘ in ‚problem črne

škatile') ukvarja s težavo, da domnevno nimamo ustreznega vpogleda v postopke odločanja (nekaterih) algoritmov in da podrobnosti vsebine, izračunov in postopkov algoritmov UI človek ne more smiselno razumeti.¹ Algoritmi prejmejo določeno količino vhodnih podatkov, ki jih nato uporabijo za produkcijo izhodnih podatkov, tj. določene konkretne klasifikacije ali odločitve. Kompleksnejši algoritmi so netransparentni v smislu, da ko prejmemo njihov rezultat (npr. določeno odločitev o klasifikaciji), ne vemo, kako in zakaj je bila določena klasifikacija oblikovana ali izbrana (in ali je bila v sam postopek morda vključena pristranskost) (Zhou in Chen 2018). Gre za problem transparentnosti in vprašanja, povezana z razločljivostjo in pristno motnostjo – v nasprotju z motnostjo, ki je posledica ‚tajnosti‘ (npr. države ali podjetja ne razkrivajo svojih algoritmičnih modelov, tj. algoritmov, obsega njihove uporabe in/ali uporabljenih podatkov), ali motnostjo, ki je posledica ‚tehno-loške nepismenosti‘, tj. da nekateri načel in delovanja obravnavanih sistemov odločanja ne razumejo (Burrell 2016). Bistvo te oblike netransparentnosti izhaja iz dejstva, da algoritme odločanja (UI in SU) pogosto težko razumejo celo strokovnjaki, ki jih programirajo (Mittelstadt et al. 2016), in da tisto, kar te težave povzroča, ni le dolžina kode in število ljudi, ki kodo pišejo, ampak tudi nenehne spremembe v logiki postopka odločanja algoritma – ta se namreč uči iz učnih podatkov, zaradi česar sta sledenje in interpretacija (vključno z razlago) izjemno težko dosegljiva, če ne celo nemogoča.

3. Etični problemi, ki jih (ne)transparentnost odpira

Netransparentnost povzroča vrsto problemov, ki so tako praktične (tehnične) kot etične narave. Problemi segajo od ‚zaznavanja napak, pristranskosti in diskriminacije‘, pa vse do vprašanj ‚odgovornosti‘ in ‚zaupanja‘. Večina problemov je med seboj tesno povezanih in tako drug na drugega močno vplivajo. Prvi očiten praktični (tehnični) problem je, da netransparentnost delovanja onemogoča kvalitetno zaznavo napak v algoritmu in njegovem odločanju. To pomeni, da je pri nalogah, ki jih algoritem opravlja, in nimajo jasnega (preverljivega) pravnega odgovora (npr. odobritev kredita; izbira kandidata za novo delovno mesto), napako v algoritemskem odločanju pogosto skoraj nemogoče odkriti. Veliko časa lahko mine do ugotovitve, da algoritem na primer pripisuje preveliko (glede na zgornja dva primera) težu podatku o rasi osebe, o kateri odloča – praktični problem težavnega zaznavanja napake pripelje do etičnega problema potencialne pristranskosti oz. diskriminacije. Ena izmed ključnih prednosti algoritemskega odločanja (v primerjavi s človeškim) naj bi bila ‚popolna objektivnost‘, ‚nezmožnost biti pristranski‘, ta pa se je, kot smo lahko videli, hitro izkazala za vprašljivo (Mittelstadt et al. 2016). Težava je še v tem, da tudi ko vemo, da je algoritem naredil napako (npr. na fotografiji rume-

¹ Opomba o terminologiji: običajno uporabljamo izraz algoritem kot krajši in enostavnejši način za sklicavanje na niz, ki vključuje tudi elemente, kot so modeli odločanja, priporočila ali druge posebne uporabe algoritmov za določeno nalogo (vključno z izvajanjem in uporabo tehnologij, ki na njih temeljijo), saj je taka uporaba razširjena v literaturi o obravnavani temi. Kadar so te razlike pomembne, to izpostavljamo.

ne žoge je prepoznal luno), sta odpravljanje napak in optimizacija algoritma pogosto težki nalogi, saj sta mesto in vzrok napake pogosto le težko (iz)sledljiva. Tu se skriva tudi vir enega glavnih etičnih problemov netransparentnih sistemov – vprašanje odgovornosti. Ker je mesto in vzrok napake pogosto težko odkriti, je težko določiti tudi, kdo naj za posledice zaznane napake odgovarja: programer algoritma; podjetje, ki algoritem ponuja; podjetje, ki algoritem uporablja; algoritem sam, ker se tudi uči sam ... (Coeckelbergh 2019). Vprašanje odgovornosti je nenazadnje tesno povezano še s problemom zaupanja – sistemom, pri katerih je napako, njeno mesto in vzrok težko izslediti ter posledično težko jasno določiti, kdo za potencialne napake odgovarja, le stežka zaupamo. Problem zaupanja nima le praktičnih posledic, temveč odpira tudi etična vprašanja, npr. ali nam pripada pravica, da algoritemsko odločanje o naši ‚usodi‘ (npr. pri zaposlitvi na novem delovnem mestu) zavrnemo, če sistemom ne zaupamo? In ali podjetju pripada pravica, da izbor zaposlenega opravi po metodi, ki ji samo pač zaupa (Felzmann et al. 2019)? Problem je resen, saj se zdi, da po eni strani veliko besede pri tem, kdaj bodo takšni sistemi uporabljeni za odločanje o nas, nimamo, po drugi strani pa našo besedo zahodna etika skoraj zahteva. »Sestavni del tega, da je nekdo spoštovan kot avtonomni oblikovalec svojega življenja, je, da je vprašan za pristanek glede stvari, ki bistveno in nepovratno, nepopravljivo določajo njegovo življenje« (2019, 640), pravi Žalec in dodaja: »Bistveno poseganje v življenje osebe brez njenega pristanka pa je v nasprotju z moralno slovnico sodobne zahodne družbe.« (640) Poleg tega se v zadnjem času, ko ima algoritemsko odločanje v družbi vse večji dejanski vpliv, vse pogosteje odpirajo tudi vprašanja, povezana z zlorabo netransparentnih sistemov. Pojavlja se namreč zaskrbljenost, da bi podjetja, posamezniki, institucije, ki algoritemske sisteme upravljajo, lahko prirejali rezultate oz. odločitve svojih algoritmov sebi v prid, saj je preverljivost procesa odločanja (in tako verodostojnost rezultatov) zaradi netransparentnosti delovanja pogosto nemogoča. S formulacijo ‚ne vemo sicer zakaj, a algoritemski sistem X, ki mu zaupamo, nam je jasno povedal, da Y‘ je namreč mogoče upravičiti celo vrsto spornih ‚odločitev‘ Y (do katerih morda nikoli sploh ni prišlo). Z vidika etike je glavni raziskovalni izziv, kako povezati transparentnost in možnosti uporabe UI s koncepti zaupanja, odgovornosti in dolžnosti oz. obveznosti (ki so bili prvotno oblikovani v povezavi s človeškim delovanjem). Gre za povezavo med splošnim vprašanjem glede zaupanja – vrste zaupanja, pozitivni in negativni dejavniki zaupanja – in odgovornosti s posebnimi vprašanji, ki se pojavljajo ob uporabi avtomatiziranega odločanja. Z naraščanjem uporabe avtomatiziranih sistemov je pomen tega raziskovalnega problema zelo pomemben in v javni prostor prinaša povsem nove razsežnosti – začenši s tem, da takšni sistemi postajajo zelo učinkoviti in produktivni, kar je pogosto kriterij, ki lahko premaga druge vidike, kot so pravice posameznikov, krivica in nepoštenost, možna povzročena škoda itd. Glavne teorije zaupanja poudarjajo, da sta manjša transparentnost in večja negotovost negativna dejavnika zaupanja, kar vključuje tudi zaupanje v nove tehnologije in subjekte, ki takšne tehnologije v določenem kontekstu uporabljajo. Težava transparentnosti UI zahteva ponovno konceptualizacijo nekaterih vidikov zaupanja. Algoritmi UI lahko vključujejo manipulacijo z veliko količino osebnih in

drugih podatkov in/ali sprejemajo odločitve, ki močno vplivajo na življenje posameznikov – če so nepregledni (netransparentni), je močno oteženo odkrivanje napak (ki so npr. povzročile določeno škodo) ter ustrezno prepoznavanje in določanje, kdo je (pravno) odgovoren za povzročeno škodo (Mittelstadt et al. 2016). Ker nove tehnologije in UI na družbo in posledično na vse ljudi vplivajo vedno bolj, mnogi menijo, da bo treba vprašanja, od katerih je odvisna tudi ustrezna pravna ureditev, ki bo zamejila uporabo sistemov UI na področjih, na katerih težko predvidimo vse možne posledice njihove uporabe in zahtevajo previdnost (Globokar 2019, 620–621), reševati po poti družbene razprave, v katero bodo lahko vključeni vsi (Žalec 2019, 640).

Zakaj je problem transparentnosti še pomemben? S splošne perspektive ter perspektive etike sta presojanje in odločanje v središču naših praks vrednotenja in sta vključena v ključne družbene institucije. Transparentnost, razločljivost, razumljivost, sledljivost, preglednost, poštenost, nediskriminatornost, predvidljivost, odgovornost itd. so vse merila in družbene zahteve oz. načela, ki veljajo za ljudi ali ustanove v družbenih funkcijah. Ko algoritmi UI prevzamejo kognitivno delo s socialnimi dimenzijami in kognitivne naloge, ki so jih prej izvajali ljudje ali institucije, prevzamejo družbene zahteve, vključno z zgoraj omenjenimi merili (Bostrom in Yudkowsky 2014). Ker se uporaba takšnih algoritemskih sistemov hitro povečuje (obdelava medicinskih in zdravstvenih podatkov, analiza strank in oglaševanje, znanstvene aplikacije, upravljanje s človeškimi viri, varnostne aplikacije, odkrivanje in preprečevanje kriminala, razvoj smrtonosnih avtonomnih orožij itd.), se pojavlja izziv, kako zgoraj omenjena merila izvajati in pregledovati (Yeung in Lodge 2019) – in ta izziv postaja iz dneva v dan večji. Ta splošni vidik se odraža tudi v najnovejših in nastajajočih dokumentih EU o urejanju in usmerjanju politik (European Group on Ethics in Science and New Technologies 2014; 2018), saj je transparentnost vključena med sedem ključnih zahtev za oblikovanje in izvajanje zaupanja vredne umetne inteligence (hkrati z vidiki zakonite, etične in robustne UI – transparentnost je vključena v vse tri vidike). Eden izmed raziskovalnih problemov (hkrati pa tudi družbenih izzivov) je razumeti, kaj zahteva po transparentnosti dejansko pomeni in kako oblikovati resnično razločljive sisteme UI. Pogosto deklariran cilj v zvezi s prakso UI je ‚zaupanja vredna UI‘. Ključna sestavina zaupanja vredne UI je sposobnost sistema, da svoje odločitve pojasni – sistem mora biti sposoben razložiti svoje odločitve in svoje razloge na človeku razumljiv način. To zahteva sposobnost sistema, da samodejno oceni, kaj je ljudem razumljivo in kaj ne.

4. Primeri rabe obravnavanih tehnologij v izrednih razmerah oz. v nedavni pandemiji

V okviru nedavne pandemije zaradi virusa SARS-CoV-2 so se modeli in sistemi UI široko uporabljali tako za identifikacijo obolelih kot tudi za napovedovanje lokacije in možnosti okužb, hkrati pa so služili tudi na sekundarni ravni – za oblikovanje vzorcev pri odzivih na pandemijo. Uporabo sistemov UI v pandemiji bi lahko pri-

kazali s pregledom primerov sistemov v vseh ključnih korakih: ‚detekcija, preventivni ukrepi, reakcija/odziv na virus, soočanje s posledicami in pomoč pri znanstvenem raziskovanju‘.

Na področju ‚detekcije‘ se umetna inteligenca uspešno uporablja predvsem za ‚zgodnje prepoznavanje grožnje in postavljanje diagnoze‘. Najprej si nekoliko podrobneje oglejmo primer sistema za zaznavo prihajajočih nevarnosti. Nekateri algoritmi relativno hitro in uspešno zaznavajo prihajajočo grožnjo s pomočjo prepoznavanja vzorcev iz zbranih podatkov, na podlagi katerih jim nato uspe zaznati specifične anomalije in t. i. digitalne ‚dimne signale‘, ki kažejo na bližajočo se nevarnost. Dober primer takega algoritma je BlueDot. BlueDot je program, ki so ga v Torontu začeli razvijati po izbruhu SARS-a leta 2003. Namenjen je odkrivanju, spremljanju in napovedovanju širjenja nalezljivih bolezni. Deluje tako, da zbira podatke o več kot 150 boleznih po vsem svetu – podatke zbira vsakih 15 min, 24 ur na dan. Sicer uporablja tudi podatke Centra za nadzor bolezni in Svetovne zdravstvene organizacije, a njegovi največja prednost in moč sta, da upošteva tudi manj strukturirane informacije. Večina BlueDotovih sposobnosti napovedovanja izhaja namreč ravno iz podatkov, ki niso del uradnih zdravstvenih virov. Spremlja na primer gibanje štirih milijard potnikov, ki letno potujejo s komercialnimi letali; spremlja tudi mnoge podatke, širše vezane na človeško in živalsko populacijo (tj. rodnost, umrljivost, spreminjanje pričakovane življenjske dobe itd.); vremenske podatke iz satelitov; lokalne novice; vsak dan pregleda tudi več kot 100000 spletnih člankov v 65 jezikih (Stieg 2020). Pri svojem napovedovanju uporablja kompleksne načine združevanja podatkov: združuje na primer podatke, v katerem mestu so zaznali novo bolezen, s podatki, v katere države prebivalci tega mesta najpogosteje potujejo in ali je na teh destinacijah za virus primerno podnebje. Razvijalci programa so podatke najprej razvrstili ročno in razvili taksonomijo, ki omogoča, da so relevantne besede v člankih skenirane učinkovito, nato pa so za urjenje programa uporabili strojno učenje in procesiranje naravnega jezika. Posledica natančnega strojnega učenja je, da BlueDot zelo malo primerov označi za takšne, ki bi potrebovali dodatno človeško analizo. Takoj ko program zazna specifične vzorce, ki jih oceni kot relevantne, o tem obvesti tako zdravstvene institucije in državne ustanove kot podjetja, ki so njegovi naročniki (Stieg 2020). Malo po polnoči 30. 12. 2019 je BlueDot v Wuhanu zaznal skupek (točno 27) primerov ‚nenavadne pljučnice‘ in jih povezal z Wuhansko tržnico z morsko hrano in živimi živalmi. Poleg obvestila je BlueDot kot ogrožena mesta navedel tudi lokacije, ki so (predvsem preko letalskega prometa) tesno povezane z Wuhanom, to pa je storil na zelo konkreten način (glede na opravljene nakupe vozovnic) in ne le glede na splošno oceno povprečja gibanja tamkajšnjega prebivalstva. V skladu s tem je na spisek rizičnih mest uspešno uvrstil mesta, ki so se nato tudi dejansko srečala s prvimi okužbami izven Wuhana (Bangkok, Hongkong, Tokio, Tajpej, Phuket, Seul, Singapur). Vse to je BlueDot opravil že 9 dni pred trenutkom, ko je Svetovna zdravstvena organizacija javno izjavila, da so zasledili pojavitev novega virusa. Cilj takšnih sistemov je, kot pravi Kamran Khan (ustanovitelj in direktor BlueDota), »širiti znanje hitreje, kot se širijo bolezni« (Stieg 2020). BlueDot je načeloma svojo

nalogo opravil zelo dobro (napoved je bila hitra in pravilna), a zaradi pomanjkanja zaupanja v tovrstne sisteme je bil domet vpliva njegove napovedi seveda močno omejen. Verjetno smo še daleč od tega, da bi sprožali konkretne družbene spremembe (zaprtje letališč, omejevanje gibanja in zbiranja) zgolj na podlagi algoritemske napovedi.

Popolno zaupanje tem sistemom seveda razumljivo in upravičeno še ni mogoče – in k temu prispeva kar nekaj problemov: problem prevzemanja odgovornosti za neupravičene ekonomske posege (npr. zaprtje letališč), ki so posledica napačne napovedi (težko zaupamo, če ne vemo, kdo za napoved odgovarja), in težka preverljivost uspeha napovedi (če so napoved epidemije in ukrepi pravočasni in uspešni, do epidemije načeloma sploh ne pride). Gotovo pa je zelo pomemben del problema zaupanja in s tem pomanjkanja dejanskega vpliva, ki bi ga napovedovalni sistemi, temelječi na algoritemskem odločanju, lahko imeli, tudi pomanjkanje transparentnega delovanja. Transparentno delovanje bi nam omogočilo vpogled v odločitvene procese, ki jih je algoritem opravil, razumevanje in sledljivost algoritemskega delovanja pa bi naše zaupanje upravičila veliko lažje. Zdi se torej, da bi povečana transparentnost algoritemskih odločitvenih procesov zaupanje povečala, povečano zaupanje pa bi lahko narekovalo dejansko delovanje, kar bi skrajšalo odzivni čas družbe na napovedano nevarnost – to pa je pravzaprav cilj napovedovanja nevarnosti.

Drugo relevantno polje, vezano na področje ‚detekcije‘, je ‚diagnostika‘. Uspešno in hitro diagnosticiranje bolnikov je v času pandemije ključnega pomena. Zdi se, da zamujanje simptomov virusne okužbe za kužnostjo bolnika ter širjenje virusa s takšno hitrostjo, da sredstev in osebja, ki bi pri diagnosticiranju lahko pomagali, močno primanjkuje, sama po sebi kažeta, da je uporaba UI tudi na tem področju zelo na mestu. UI se v diagnostiki COVID-19 uporablja predvsem v obliki avtomatiziranega pregledovanja CT-slik. Avtomatizirano prepoznavanje vzorcev na podlagi medicinskih slik in predhodno zbranih podatkov o simptomih močno olajšuje delo zdravnikov – ročno pregledovanje CT-slike lahko traja tudi do petnajst minut, medtem ko algoritmi UI enako (včasih natančnejšo) analizo opravijo v približno desetih sekundah (Imaging COVID-19 AI. 2020).

Na področju ‚preventivnih ukrepov‘ je pomembna uporaba UI za ‚napovedovanje‘, ‚nadzor‘ in ‚informiranje‘. Pri napovedovanju UI omogoča relativno zanesljivo napovedovanje možnosti okužbe posameznikov in smeri širjenja virusa. Kot primer navedimo EpiRisk – spletno aplikacijo (algoritem UI), ki izračunava tveganja in smer širjenja bolezni glede na strukturo topologije letalskih povezav po svetu (The GLEAM Project 2020). Z izrazom ‚nadzor‘ mislimo zlasti na opazovanje in spremljanje širjenja okužb v dejanskem času z mnogimi aplikacijami, ki npr. preko kamer spremljajo varnostno razdaljo med posamezniki na letališčih ali pa s sledenjem pametnim telefonom nadzorujejo gibanje posameznikov, ki so bili okuženi ali v stiku z okuženimi ljudmi. Algoritmi UI so v pomoč tudi pri zbiranju, filtriranju in širjenju pomembnih informacij npr. po družbenih omrežjih, kar pomaga v boju proti slabi seznanjenosti prebivalstva ali širjenju lažnih informacij. Mnoge aplikacije, ki temeljijo na sistemih UI, omogočajo personalizirane novice, uporabniku ponujene glede na njegovo lokacijo in osebne preference. Te preference so sle-

dljive iz vzorcev, ki so posledica zbranih podatkov iz osebnih iskanj posameznika po spletnih brskalnikih.

Najpogosteje omenjeno področje uporabe UI v ‚odzivu oz. reakciji je avtomatizacija storitev‘. Dober primer je trojica robotov, ki so jih razvili pri UBTECH Robotics in jih uporabljajo v bolnišnici v Shenzhenu: ATRIS, AIMBOT in Cruzr. ATRIS je robot, ki v okolici bolnišnice (tudi ponekod drugod v mestu) med vožnjo predvaja zvočna navodila za boj proti pandemiji, prepoznava obraze brez mask, ljudi na to opozarja in razkužuje zunanje površine. Vse to pogosto opravlja sam, brez dodatnega vodenja. Cruzr je različica robota, ki se vozi po bolnišnici in pomaga pacientom, ki potrebujejo kakršne koli informacije. Ponuja tudi t. i. ‚antiepidemiološke konzultacije‘: posameznikom v dialogu zagotavlja odgovore na vsa njihova vprašanja v povezavi s pandemijo in bojem proti njej. Cruzr s kamero v predelu ‚obraza‘ zdravnikom omogoča, da pacientom postavljajo diagnoze, izvajajo razgovore in posvete na daljavo. AIMBOT je tretja različica robota, ki med drugim meri telesno temperaturo ljudi, ki se sprehajajo po avli bolnišnice. Strokovnjaki pravijo, da zdravniki v njihovi bolnišnici opravijo v povprečju 6 izmen meritev temperature na dan, AIMBOT pa temperaturo izmeri kar 200 ljudem na minuto, saj jo lahko meri istočasno vsem v njegovem vidnem polju (Ackerman et al. 2020). Sam tudi razkužuje prostore bolnišnice (medicinsko osebje je to izvajalo 6-krat na dan, AIMBOT razkuževanje opravlja 24 ur na dan) in meri varnostno razdaljo med posamezniki, kršitelje pa na neustrezno opozarja.

Pri ‚soočanju s posledicami‘ se UI uporablja npr. za sledenje ekonomskemu okrevanju mnogih držav. To sledenje po večini opravljajo programi, ki se zanašajo na satelite, GPS in podatke iz družbenih omrežji. Pomoč ob ekonomskih posledicah krize nudi WeBank – digitalna banka, ki ponuja storitve (na primer osebno finančno svetovanje), temelječe na strojnem učenju in orodjih za procesiranje naravnega jezika, govora in spletnega preverjanja identitete. Orodje med drugim samo izvaja storitev osebnega finančnega svetovanja (WeBank Co., Ltd. 2014).

Morda skoraj najbolj ključen vidik uporabe umetne inteligence pa je ‚pomoč pri znanstvenem raziskovanju‘. Sistemi UI so v veliko pomoč pri odkrivanju lastnosti virusa, razumevanje katerih je ključnega pomena pri vzpostavitvi načrta za boj proti širjenju (razumeti želimo načine širjenja, odpornost virusa na temperaturo, imunito ozdravljenih ljudi itd.) ter pri odkrivanju potencialnega zdravila in cepiva. Dober primer je AlphaFold, ki so ga razvili pri družbi DeepMind. Velja omeniti, da laboratoriji z vsega sveta že od začetka pandemije svoje ugotovitve med seboj delijo odprto, saj sta transparentnost rezultatov in njihovo deljenje v kriznih časih ključnega pomena za kar se da hiter napredek pri iskanju rešitev. Sodelovanje laboratorijev se v času pandemije zdi ne le nujno, temveč skoraj samoumevno, razlog pa je verjetno v tem, da se mnogi razlogi za namerno netransparentnost rezultatov (npr. poslovna skrivnost – prednost pred konkurenco) v času pandemije izkažejo za ne najbolj prepričljive. Ravno ob takšni pomoči različnih laboratorijev se je AlphaFold lotil napovedovanja proteinske strukture virusa (Jumper et al. 2020). Algoritem uporablja modele nevronske mreže, ki omogočajo, da na podlagi mnogih podatkov išče specifične vzorce in povezave. Te vzorce se nato nauči uporabljati za iskanje

podobnih vzorcev na novih podatkih. AlphaFold uporablja ta koncept za napovedovanje lastnosti strukture proteinov, analiziranje, ki bi bilo za ljudi dolgotrajno, pa opravi mnogo hitreje (Gulamali 2020). Svojo napoved proteinske strukture SARS-CoV-2 je tako podal že kmalu po razglasitvi pandemije, a v času pisanja tega prispevka proteinska struktura še ni dokončno raziskana in znana; na oceno uspešnosti AlphaFoldove napovedi moramo še nekoliko počakati (Scudellari 2020). Poznavanje oblike proteina bi lahko omogočilo boljše razumeti funkcije v bioloških procesih, povezanih z virusom – to pa je ključno za pripravo zdravila in cepiva, ki bosta v boju proti COVID-19 uspešna. Ko in če bosta zdravilo in cepivo na voljo, bo torej vsaj del zasluga skoraj zagotovo mogoče pripisati tudi algoritemskim sistemom UI.

Algoritmom se pripisuje, da so natančnejši, hitrejši, objektivnejši in večina teh lastnosti izhaja iz dejstva, da imajo opravka z ogromnimi količinami podatkov, ki jih (zaradi svoje visoke zmogljivosti) obdelujejo hitro in natančno. Zdi se, da je največji problem uporabe sistemov UI pri pandemiji COVID-19 ravno pomanjkanje določenih podatkov. Sistemi npr. dobro napovedujejo širjenje virusa, opravljajo diagnoze in spremljajo stike med ljudi (varnostno razdaljo med njimi ipd.), saj so vsi podatki, potrebni za sledenje in napovedovanje, dostopni. Nekoliko počasnejša pa je njihova uporaba pri iskanju zdravil in cepiva, saj relevantnih podatkov (npr. o strukturi virusa) ni v izobilju. Ravno zato so znanstveniki stopili korak nazaj in sisteme UI začeli uporabljati tudi za odkrivanje lastnosti virusa, tako da je podatkov, ki bodo na voljo sistemom algoritemskega odločanja, vedno več – kot rečeno, delno zahvaljujoč ravno umetni inteligenci. Med pandemijo je zaradi vse večje količine podatkov uporaba umetne inteligence tako vse boljša. Zdi se, da se sistemi strojnega učenja ne učijo (povečujejo svojo natančnost, uspešnost in zanesljivost) le med posamezno pandemijo, pač pa je njihov razvoj in napredek očiten ob vsaki novi pandemiji. Razvoj BlueDota je sprožila ideja ob izbruhu SARS-a leta 2003; postopoma so ga vse bolj uporabljali za napovedovanje širjenja mnogih virusov (leta 2014 je napovedal, da se bo ebola razširila iz Zahodne Afrike in leta 2016, da se bo virus zika razširil na Florido). Decembra 2019 je prvi uspešno napovedal, da se bo COVID-19 razširil po vsem svetu. Pričakujemo lahko torej, da bo umetna inteligenca v prihodnjih pandemijah igrala še bolj ključno vlogo, in to ne le zaradi lastnega razvoja, temveč tudi zaradi ,trenin-ga', ki ga je algoritemskim sistemom omogočila prav pandemija COVID-19.

5. Etika izrednih razmer ter dopustnost rabe obravnavanih tehnologij

Opisali smo že, kako netransparentnost samodejnih sistemov odločanja odpira cel niz etični in drugih dilem. Pred umestitvijo rabe takšnih sistemov v okvir izrednih razmer lahko izpostavimo tudi pravne vidike. Evropska unija je trenutno vodilni akter v gibanju za transparentno, razložljivo in na človeka usmerjeno UI in je že storila več korakov v zvezi z urejanjem tega področja, a hkrati jasno priznava pomembno vlogo UI za družbo in njen razvoj. V Splošni uredbi o varstvu podatkov (GDPR; zlasti člena 22 in 13) je navedeno, da regulacija UI vključuje »pravico, da

oseba ni podvržena zgolj avtomatiziranemu odločanju, razen v določenih situacijah«. Prav tako vključuje posebne zahteve glede transparentnosti pri uporabi avtomatiziranega odločanja, in sicer obveznost obveščanja o obstoju/poteku takšnih odločitev in zagotavljanja pomenljivih informacij (pomenljivih razlag) ter pojasnjevanja njihovega pomena in predvidenih posledic za posameznika (Malgieri in Commandé 2017). Po drugi strani nekateri trdijo, da obstaja več razlogov za dvom v obstoj, obseg in izvedljivost »pravice do obrazložitve oz. razlage«, ki je povezana z avtomatiziranim odločanjem (Wacher et al. 2017; 2019). Odprto ostaja vprašanje, kako te domnevne pravice formulirati in, kar je še pomembneje, kako zamejiti in razdeliti odgovornost med različne vpletene organe (vlade, mednarodne organizacije, poslovne korporacije itd.) – to je eden osrednjih nerešenih problemov (Wachter et al. 2018). Vseeno si lahko zastavimo vprašanje, ali niso izredne razmere epidemij in pandemij dovolj tehten razlog, da lahko mnoge izmed etičnih, pravnih in tudi širše družbenih pomislekov postavimo na stran, posebej v luči izkazane učinkovitosti in koristnosti mnogih obravnavanih sistemov. V iskanju odgovora se bomo obrnili k področju etike izrednih zdravstvenih razmer.

Etika izrednih zdravstvenih razmer, kakršne predstavljajo med drugim epidemije in pandemije, sodi v okvir t. i. etike javnega zdravja. Pomembno je najprej razumeti, da slednja vključuje tako vprašanja o odzivih na takšne razmere kot tudi o pripravljenosti sistemov javnega zdravstva ter družbe nasploh na takšne dogodke. Ne gre torej zgolj za določanje etičnih okvirov delovanja, ko se izredne zdravstvene razmere pojavijo, temveč tudi za dolžnosti in druge etične premisleke, ki smo jim zavezani že prej. Izredne razmere za področje javnega zdravja pomenijo »situacijo, v kateri lahko posledice na področje zdravja povsem presežejo običajne zmožnosti družbe, da le-te naslovi« (Ellis et al. 2016, xxii). Epidemije in pandemije, podobne nedavni pandemiji virusa SARS-CoV-2, v ta okvir vsekakor sodijo – so pa njihove posledice lahko izjemno pomembne tudi za druga področja življenja. Tu nas bodo zanimali predvsem vidiki etičnih načel in vodil v takšnih situacijah, posebej tisti, ki zadevajo problem transparentnosti. Eno izmed vodil na področju etike izrednih zdravstvenih razmer je, da takšne razmere sicer predstavljajo poseben izziv za etično odločanje in presojo, vendar pa ne terjajo posebnih ali prilagojenih etičnih načel/vrednot (Ellis et al. 2016, xxii). Tudi zato je ob drugih posebej izpostavljen vidik pripravljenosti na takšne razmere (v smislu prožnosti tako družbe kot zdravstvenih, gospodarskih oz. finančnih, političnih, okoljskih in družbenih struktur, ki družbo vzpostavljajo).

Kakšno vlogo ima pri vsem naštetem transparentnost? Kot etično izhodišče, načelo in vodilo je v etiki izrednih zdravstvenih razmer prisotna predvsem v povezavi z zaupanjem (javnosti) in preglednostjo delovanja odločevalcev ter snovalcev politik. »Zaupanje javnosti in zaupanje nasploh sta ključna za pripravljenost in odzivnost na področju izrednih razmer za javno zdravje. Odločitve glede javnega zdravja bodo tako najbolj učinkovite takrat, kadar bodo transparentne in bodo neposredno povezane s skupnostmi, ki jim služijo.« (Ellis et al. 2016, xxi) Odprto ostaja vprašanje, kaj predstavlja merilo za učinkovitost odločitev oz. ukrepov. To lahko razumemo le v okviru maksimizacije zdravja ali pravične porazdelitve takšnih

najboljših učinkov prizadevanja za javno zdravje, a v nobenem primeru to ne naslavlja vprašanja transparentnosti. Pri slednjem in vprašanju uporabe sistemov UI moramo biti pozorni tudi, da gre pri razmislekih o javnem zdravju in izrednih razmerah skoraj vedno za tehtanje določenih temeljnih vrednot (varnost, zdravje, dobrobit idr.) z drugimi etičnimi premisleki (npr. kratkoročni interesi nasproti dolgoročnim interesom) in pravnimi okviri (človekove pravice, svoboda, lastnina). Treba je biti pozoren še na upravičenje oz. utemeljitev odločitev.

»Etično upravičenje je prav tako zelo pomembno za ukrepe na področju javnega zdravja, kajti večinoma se morajo vodilni na področju javnega zdravja in javne varnosti zanašati na prostovoljno upoštevanje le-teh s strani izjemno velikega števila ljudi. To prostovoljno obnašanje pa je po drugi strani odvisno od tega, ali ti ljudje prepoznavajo dobre razloge za upoštevanje oz. spoštovanje ukrepov, vključno z dobrimi etičnimi razlogi.« (Jennings in Arras 2016, 5)

Vsaj nekateri od predstavljenih sistemov odločanja UI imajo pri navedenem najmanj dve težavi. Prva zadeva prostovoljnost in obveščenost, kajti ni povsem jasno, ali gre pri teh sistemih za kakršno koli privolitev posameznikov. Druga, pomembnejša, zadeva razloge za sprejete odločitve. Netransparentnost pomeni prav to, da ne moremo izpostaviti razlogov, zakaj je neki sistem določeno odločitev sprejel oz. priporočil. Zamislimo si sistem, ki napoveduje širjenje epidemije na podlagi vzorcev in parametrov, ki nam niso razumljivi in dosegljivi. Recimo, da predvidi povišano verjetnost širjenja za regijo A v določeni državi, za regiji B in C pa tega ne predvidi, čeprav vsi običajni statistični podatki relevantnih razlik ne kažejo. Ali lahko v tem primeru rečemo, da poznamo razloge za uvedbo omejitvenih ukrepov v regiji A in da jih lahko sprejmemo kot dobre razloge (razen tega, da sistem deluje dobro)? Verjetno bi temu težko pritrdili.

Jennings in Arras (2016, 13–14) kot jedro etike javnega zdravja in izrednih razmer izpostavljata sedem temeljnih usmerjevalnih načel: zmanjševanje škode in povečevanje koristi, enakopravnost pri svoboščinah in človekovih pravicah, razdelilna pravičnost, transparentnost in vključevanje javnosti, prožnost družb in opolnomočenje, profesionalizem javnega zdravja in državljanska odgovornost ter zavzetost za skupno dobro. Transparentnost je pomembna, ker posamezniku omogoča, da lažje razume, zakaj so mu bile ob izrednih razmerah, kot je epidemija, omejene nekatere svoboščine ali naložene dodatne dolžnosti; da jih lahko razume in sprejme kot pravične in razumne. Ravno zato je poleg samih odločitev v obravnavanem kontekstu pomembno, kako odločitve oblikujemo in sprejmemo.

»Vsi ljudje so prepričani, da je njihovo življenje enako vredno kot življenja drugih ljudi, tako da če bi katera koli tragična odločitev dala prednost tem drugim pred nami ali našimi bližnjimi (na primer, da bi šlo za odločitev o tem, za koga uporabiti medicinski ventilator in ali dati cepivo nekomu drugemu ter bi zaradi tega zelo verjetno umrli ali utrpeli drugo škodo), potem bi vsekakor vztrajali, da želimo izvedeti, kdo je sprejel to odločitev in na podlagi

katerih razlogov jo je sprejel. Vsekakor bomo v takšni situaciji terjali zagotovila, da je bila odločitev pravična in dosežena v poštenem procesu.« (33)

Slednje pa terja troje: javnost ali transparentnost odločanja, možnost pritožbe oz. preveritve odločitve ter relevantnost razlogov za odločitve.

Načelo transparentnosti glede temeljev za odločitve zahteva izkazovanje transparentnosti in javnosti. To preprečuje morebitno pristranskost, predsodke, nepravičnost in napake, za katere sicer ne bi nihče nosil odgovornosti – če bi bile odločitve sprejete v tajnosti. Po drugi strani omogoča, da so odločitve podprte z dobro informiranostjo in bodo zaradi tega tudi pravilnejše (K temu prvemu vidiku se včasih dodaja še, da mora odločevalec transparentno komunicirati tudi navzven, torej tudi z javnostmi, ki jih sicer odločitev ne zadeva neposredno). Načelo možnosti pritožbe terja možnost pritožbe na odločitev ter dolžnost odločevalca, da ostane odprt za morebitne nove poglede in nove argumente proti začetno sprejeti odločitvi. Načelo relevantnosti pa terja, da naj bodo razlogi za odločitev omejeni s tem, kar lahko v družbi sprejmemo kot obče (33–35). Vidimo lahko, da je širši problem transparentnosti vpet v vsa tri omenjena načela, kajti če nimamo pravega vpogleda v sisteme UI, potem ne moremo (v celoti) zadostiti nobenemu izmed omenjenih načel. Ob povedanem lahko oblikujemo širšo in določnejšo obliko načela transparentnosti pri odločanju. Načelo transparentnosti torej terja, da morajo biti odločitve sprejete na odprt način, razlogi za odločitev morajo biti izpostavljeni in pojasnjeni, hkrati pa morajo biti takšni, da jih tisti, ki ga odločitev zadeva, lahko razume in (na načelni ravni) sprejme. Problem pri uporabi tega načela za problem transparentnosti (kot ga odpira uporaba UI in SU) je, da je oblikovano pretežno za sprejemanje odločitev pri človeških odločevalcih – in je tako primarno usmerjeno na področje pojasnjevanja postopkov sprejemanja odločitev ter njihove javnosti. Zato se v mnogočem sploh ne sooča s pravim problemom, ki ga obravnavane tehnologije prinašajo.

Glede problema transparentnosti sistemov UI bi možen premislek šel v smeri, da izredne razmere pri splošnih etičnih načelih – torej tudi načelu transparentnosti – dopuščajo izjeme.

»Vseeno ne moremo povsem kategorično izključiti prisile in zaupnosti oz. namernega skrivanja podatkov pred javnostjo, čeprav se ju moramo kar sedita izogibati. Upravičenje za takšno določitev v posamičnem primeru je stvar konteksta in okoliščin. Prisilne izselitve ali karantena so včasih neizogibne in etično upravičene. Skrivanje informacij pred javnostjo je morda kdaj nujno, da ne prihaja do panike in drugih ravnanj, ki bi v celoti stvari še mnogo bolj poslabšala.« (8)

Ta premislek bi lahko deloma upravičil uporabo obravnavanih sistemov UI, vendar vsekakor ne bi dopuščal rabe vseh. Hkrati pa bi vseeno ostajala etična obveza upravičiti in pojasniti njihovo uporabo (tudi tisto izjemno) pred izrednimi razmerami in po njih – toda zopet bi naleteli na problem transparentnosti, saj javnostim delovanja teh sistemov ne bi zmogli pojasniti v celoti. Težava za takšen premislek se kaže tudi v tem, da so ti sistemi učinkoviti že pred izrednimi razmerami, npr. pri njihovem napovedovanju.

6. Sklep

Kaj torej lahko sklenemo glede problema transparentnosti pri uporabi sistemov odločanja na podlagi UI? Na voljo so tri poti oz. odzivi. Prvi – da prilagodimo tradicionalno pojmovanje transparentnosti pri odločanju (posebej za področje zdravja) tako, da lahko tudi omenjene sisteme (ali vsaj njihovo veliko večino) razumemo kot transparentne. Zagovorniki takšnega odgovora v ospredje postavljajo predvsem dva premisleka. Prvi je, da je za transparentnost povsem dovolj, če delovanje takšnih algoritmov (predvsem s tehničnega vidika) razumemo načelno in nadzorujemo njihovo učinkovitost (da ne prihaja do znatnega deleža napak). Drugi premislek je, da tudi človeško odločanje pogosto ali celo pretežno ni zares transparentno. Zakaj bi torej takšno transparentnost terjali od sistemov algoritemskega odločanja? Drugi – prilagoditev normativnih izhodišč za to področje. Jedro tega pristopa je v zamejitvi domene relevantnosti običajnih (bio)etičnih načel in vodil tako, da obravnavane sisteme iz njih izključimo ali pa se enostavno načelu transparentnosti v celoti odpovemo. Tretji – omejitev uporabe algoritemskih sistemov odločanja zgolj na tiste sisteme, ki lahko zadostijo običajnemu oziroma tradicionalnemu pojmu transparentnosti. To bi pomenilo tudi, da bi se morali odpovedati uporabi nekaterih izmed najbolj učinkovitih orodij, ki omogočajo spopadanje z izrednimi zdravstvenimi razmerami.

Izbira med predlaganimi možnostmi je v današnjem času, ko ima UI vse večjo vlogo, nujna, a zahteva tehten premislek. Premislek, ki bo lahko nato narekoval izbiro poti, bo po našem mnenju moral vključevati vnovičen premislek o izhodiščih etične dileme ob uporabi netransparentnih sistemov odločanja. To najprej zajema premislek o samem ‚konceptu transparentnosti‘ – jasna opredelitev koncepta bo omogočila enotno razumevanje, ki je pogoj za možnost diskurza. Posebej bo treba podrobneje proučiti ‚transparentnost odločanja‘ (tako v algoritemskih sistemih kot pri ljudeh), saj je to nepogrešljiv korak pri oceni ustreznosti vzpostavitve transparentnosti odločanja kot splošnega etičnega načela – in njegove potencialne fleksibilnosti v primeru izrednih razmer. Tak premislek bo morda pokazal, ali etično načelo transparentnosti (vsaj deloma) ustvarja neutemeljena pričakovanja in morda celo nekaj lažnih dilem. Morda bo odprl tudi nekaj novih etičnih vprašanj in dilem, ki jih sedaj še niti ne slutimo. Prav gotovo pa bo pomembno prispeval k polnjenju vrzeli, ki jo manko zavezujočih in učinkovitih smernic, ki bi urejale to področje, pušča med teorijo in prakso etične uporabe umetne inteligence.

Kratice

- UI – umetna inteligenca.
- SU – strojno učenje.

Reference

- Ackerman, Evan, Eric Guizzo in Fan Shi.** 2020. Robots Help Keep Medical Staff Safe at COVID-19 Hospital. *IEEE Spectrum*, 20. 3. <https://spectrum.ieee.org/automaton/robotics/robotics-hardware/video-friday-ubtech-robots-covid-19-shenzhen-hospital> (pridobljeno 13. 7. 2020).
- Bostrom, Nick, in Eliezer Yudkowsky.** 2014. The ethics of artificial intelligence. V: Keith Frankish in William Ramsey, ur. *Cambridge Handbook of Artificial Intelligence*, 316–334. New York: Cambridge University Press.
- Burrell, Jenna.** 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* 3, št. 1:1–12.
- Coeckelbergh, Mark.** 2019. Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability. *Science and Engineering Ethics* 26, 24. 10. <https://link.springer.com/article/10.1007%2Fs11948-019-00146-8> (pridobljeno 16. 7. 2020).
- Eliazat, Adolfo.** 2020. Artificial Intelligence: The Holy See Pleads for an Ethical Algorithm. Adolfo Eliazat, 24. 4. <https://adolfoeliazat.com/2020/04/24/artificial-intelligence-the-holy-see-pleads-for-an-ethical-algorithm/> (pridobljeno 15. 7. 2020).
- Ellis, Barbara A., Drue H. Barrett, John D. Arras in Bruce Jennings.** 2016. Introduction. V: Bruce Jennings, John D. Arras, Drue H. Barrett in Barbara A. Ellis, ur. *Emergency Ethics Public Health Preparedness and Response*, xvii–xxxvii. New York: Oxford University Press.
- European Group on Ethics in Science and New Technologies.** 2014. *Opinion No. 28: Ethics of Security and Surveillance*. Bruselj: Directorate-General for Research and Innovation.
- . 2018. *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*. Bruselj: Directorate-General for Research and Innovation.
- Felzmann, Heike, Eduard Fosch Villaronga, Christoph Lutz in Aurelia Tamo-Larrieux.** 2019. Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society* 6, št. 1 (januar–julij). <https://journals.sagepub.com/doi/pdf/10.1177/2053951719860542> (pridobljeno 12. 7. 2020).
- Globokar, Roman.** 2019. Normativnost človeške narave v času biotehnoškega izpopolnjevanja človeka. *Bogoslovni vestnik* 79, št. 3:611–628.
- Gulamali, Faris.** 2020. AlphaFold Algorithm Predicts COVID-19 Protein Structures. *InfoQ*, 31. 3. <https://www.infoq.com/news/2020/03/deepmind-covid-19/> (pridobljeno 12. 7. 2020).
- Imaging COVID-19 AI.** 2020. Home page. <https://imagingcovid19ai.eu/> (Pridobljeno 12. 7. 2020).
- Jennings, Bruce, in John D. Arras.** 2016. *Ethical Aspects of Public Health Emergency Preparedness and Response*. V: Bruce Jennings, John D. Arras, Drue H. Barrett in Barbara A. Ellis, ur. *Emergency Ethics Public Health Preparedness and Response*, 1–103. New York: Oxford University Press.
- Jumper, John, Kathryn Tunyasuvunakool, Pushmeet Kohli, Demis Hassabis in AlphaFold Team.** 2020. Computational predictions of protein structures associated with COVID-19. DeepMind: Version 2, 8. 4. <https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19> (pridobljeno 11. julija 2020).
- Malgieri, Gianclaudio, in Giovanni Comandé.** 2017. Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. *International Data Privacy Law* 7, št. 4:243–265. doi: 10.1093/idpl/ix019
- Mittelstadt, Daniel Brent, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter in Luciano Floridi.** 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3, št. 2:1–21.
- Petkovšek, Robert.** 2019. Teologija pred izzivi sodobne antropološke krize: preambula apostolake konstitucije Veritatis gaudium. *Bogoslovni vestnik* 79, št. 1:17–31.
- Scudellari, Megan.** 2020. The sprint to solve coronavirus protein structures — and disarm them with drugs. *Nature*, št. 581. 19. 5. <https://www.nature.com/articles/d41586-020-01444-z> (pridobljeno 10. 7. 2020).
- Stieg, Cory.** 2020. How this Canadian start-up spotted coronavirus before everyone else knew about it. *Make It*, 3. 3. <https://www.cnbc.com/2020/03/03/bluedot-used-artificial-intelligence-to-predict-coronavirus-spread.html> (pridobljeno 10. 7. 2020).
- The GLEAM Project.** 2020. Vision. <https://www.gleamproject.org/vision> (pridobljeno 12. 7. 2020).
- Wachter, Sandra, Brent Mittelstadt in Luciano Floridi.** 2016. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law* 7, št. 2:76–99.
- WeBank Co., Ltd.** 2014. Home page. <https://ai.webank.com/pages/en.html> (pridobljeno 13. 7. 2020).
- Zhou, Jianlong, in Fang Chen, ur.** 2018. *Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent*. Berlin: Springer.
- Žalec, Bojan.** 2019. Liberalna evgenika kot uničevalka temeljev morale: Habermasova kritika. *Bogoslovni vestnik* 79, št. 3:629–642.